Protein Functionality in Food Systems

Navam S. Hettiarachchy
Department of Food Science
University of Arkansas
Fayetteville, Arkansas

Gregory R. Ziegler
Department of Food Science
Pennsylvania State University
University Park, Pennsylvania

2
Solubility of Proteins: Protein–Salt-Water Interactions

Thomas F. Kumosinski and Harold M. Farrell, Jr.

Agricultural Research Service
United States Department of Agriculture
Philadelphia, Pennsylvania

INTRODUCTION

Biotechnology holds the promise of developing new designer-type products for food science with tailor-made functionalities via genetic engineering of proteins and creation of new co-solutes, which may control functionality through chemical, biochemical, or genetic techniques (Richardson et al., 1992). However, the historic problem of developing quantitative measures for structure-function relationships still plagues the researcher. Without knowledge of these relationships, the new techniques are limited to costly hit-or-miss experiments, which have a low probability of success.

The caseins of bovine milk and their naturally occurring genetic variants provide an illustration of qualitative correlations between primary structure and protein functionality in food systems. For example, milks

Reference to a brand or firm name does not constitute an endorsement by the U.S. Department of Agriculture over others of a similar nature not mentioned.

containing α_{s1}-casein A rather than the more frequently occurring α_{s1}-casein B variant yield cheeses with a softer texture and body; at the same time these milks are more resistant to calcium-induced coagulation, i.e., they are more stable at elevated calcium concentrations (Thompson et al., 1969). Here the A variant is the result of the sequential deletion of 13 amino acids (residues 14 to 26) from the B variant. However, changes in protein secondary, tertiary, and quaternary structure as well as in the thermodynamic parameters resulting from this mutation have not been obvious. Thus, a complete thermodynamic and structural mechanism for the above functionality changes has not been elucidated, and the success rate of future technologically induced mutations cannot be predicted.

Finally, in recent years, the emergence of molecular modeling as a technique for refining existing three-dimensional molecular structures or building new predicted models has yielded a methodology with the capability of developing a molecular basis for structure-function relationships (Kumosinski et al., 1991a,b). Now, not only food proteins but preservatives, salts, stabilizers, etc., may be modeled for their potential effectiveness in structure-function relationships.

We have attempted to define and model one simple functionality test for the caseins: solubility as a function of calcium ion concentration. This system was selected for three reasons. From the point of view of the food industry, caseinate is an important commodity and milk and dairy products are widely consumed for their calcium content. Second, the interactions occurring in this important colloidal-transport system are still not well defined. Third, a wealth of information of a qualitative nature is available in the literature on calcium-induced casein solubility curves (Arakawa and Timasheff, 1984; Farrell and Kumosinski, 1988; Farrell et al., 1988; Kumosinski and Farrell, 1991). In order to better understand these calcium-protein interactions, the precipitation and resolubilization of selected caseins were reinvestigated. The functionality data were analyzed with respect to computer-generated models; analysis of the data indicates that a thermodynamic linkage occurs between calcium binding and salting-out and salting-in reactions.

Thus, a quantitative thermodynamic mechanism could be established for the salting-in and salting-out of casein, and binding free energies for these ligand-induced protein solubility profiles may easily be calculated. Finally, molecular modeling techniques such as energy minimization and molecular dynamics were utilized to mimic protein–salt-water interactions, which explain the salt-induced solubility profiles of α_{s1}-casein. Here, a recently developed predicted energy-minimized three-dimensional structure of α_{s1}-casein was employed to discover the hydrophobic

sites responsible for precipitation of the protein and potential salt-binding sites responsible for the salting-in process with added divalent salts.

THEORY

Thermodynamic Linkage

Wyman's theory of thermodynamic linkage (Wyman, 1964) is based on the concept that changes in an observable physical quantity (in this case solubility) can be linked to ligand binding. In previous studies on isolated caseins (Farrell et al., 1988; Kumosinski and Farrell, 1991) it was shown that the precipitation of the caseins in the presence of calcium is indeed linked to calcium binding, and that calcium binding is the driving force in both salting-out and salting-in.

Here, we assume that there are essentially two classes of binding sites for ligands responsible for the sequential salting-out and salting-in processes and, therefore, Wyman's linked functions equations (1964) can be used to treat these processes with the assumption that the following equilibria occur:

$$\underset{(S_0)}{p} + nX \overset{k_1^n}{\leftrightarrows} \underset{(S_1)}{PX_n} + mX \overset{k_2^m}{\leftrightarrows} \underset{(S_2)}{PX_nX_m} \tag{1}$$

where p is the unbound protein, X is the free salt, n and m are the number of X moles bound to species PX_n and PX_nX_m, and S_0, S_1 and S_2 are the solubilities of the species indicated. For this study S_1 and S_2 will be relative to S_0. The mathematical relationship representing the above stoichiometry can be represented according to the following:

$$S_{app} = S_0 f(p) + S_1 f(PX_n) + S_2 f(PX_nX_m) \tag{2}$$

where S_{app} is the apparent protein solubility at a given salt concentration (X_T), $f(i)$ are the protein fractional component of species i and the S's are species previously defined. Incorporation of the salt-binding equilibrium constants (k_1 and k_2) as defined by Eq. (1) into Eq. (2) yields the following:

$$S_{app} = \frac{S_0 p}{p + k_1{}^n p x^n} + \frac{S_1 k_1{}^n p x^n}{p + k_1{}^n p x^n} + \frac{(S_2 - S_1) k_2{}^m p x^m}{p + k_2{}^m p x^m} \tag{3}$$

where p is the concentration in percent of the unbound protein and x is the concentration of unbound salt. Cancelation of common terms yields:

$$S_{app} = \frac{S_0}{1 + k_1{}^n x^n} + \frac{S_1 k_1{}^n x^n}{1 + k_1{}^n x^n} + \frac{(S_2 - S_1) k_2{}^m x^m}{1 + k_2{}^m x^m} \tag{4}$$

It should be stressed here that the above expression is valid for sequential

binding, i.e., $k_1 > k_2$, and n sites saturate prior to the binding of m sites on the protein and, for simplicity, that n and m do not interact. Also, for n or $m > 1$, k_1 and k_2 represent an average value for each class of the n or m binding sites. In reality n or m moles of salt will bind with only one equilibrium constant (K_1), i.e., $K_1 = k^n{}_1$ and $K_2 = k^m{}_2$.

Now, since the total salt concentration, X_T, is the sum of the free salt concentration, x, and the concentration of the bound salt of both species PX_n and PX_nX_m, it can be shown that

$$X_T = x\left(1 + \frac{nk_1{}^nP_Tx^{(n-1)}}{1 + k_1{}^nx^n} + \frac{mk_2{}^mP_Tx^{(m-1)}}{1 + k_2{}^mx^m}\right) \tag{5}$$

where P_T is the total concentration of protein. From Eq. (5) it can be seen that X_T approaches x when P_T is small relative to x. In some of our experiments this assumption is reasonable. For α_{s1}-casein, very low (1–2 mM) concentrations of calcium induce relatively strong aggregations (Waugh et al., 1971). This suggests that the kinetically active unit in precipitation is an aggregate, so that the number of moles of aggregate is smaller than the number of moles of total protein. In direct protein solubility studies this may not be the case. Using the values of K_a published by Dickson and Perkins (1971) for α_{s1}-casein along with P_T, free x can be calculated and used to generate adjusted k_i's which may be more properly related to an "apparent" binding constant (Kumosinski and Farrell, 1991). However, the latter is still an approximation, and from the point of view of the kinetically active species, X_T is also appropriate.

It could be argued that the calcium-induced protein self-associations are a complicating factor in these studies even though the monomer and aggregate are both soluble. Cann and Hinman (1976) developed equations for dealing with the effects of association on ligand binding. Using their concepts, it can be seen that:

$$\begin{matrix} S_0 & k^j{}_0 & S_0 & k^n{}_1 & S_1 \\ kp + jx & \rightleftharpoons & P_kX_j + nx & \rightleftharpoons & P_kX_fX_n \end{matrix} \tag{6}$$

and the apparent solubility S_{app} again is equal to the sum of the solubilities of each species times the fraction of the protein in that species.

$$S_{app} = S_0fp + S_0\,(fP_kX_j - f\,P_kX_fX_n) + S_1fP_kX_fX_n \tag{7}$$

Thus:

$$S_{app} = S_0fp + S_0fP_kX_j + (S_1 - S_0)\,fP_kX_fX_n$$

$$S_{app} = S_0\left[\frac{p}{p + k^j{}_0P^kx^j} + \frac{k^j{}_0P^kx^j}{p + k^j{}_0P^kx^j}\right] + (S_1 - S_0)\,\frac{P_kX_fX_n}{P_kX_j + P_kX_fX_n} \tag{8}$$

and

$$S_{app} = S_0 \frac{k'_0 P^k x^j + p}{p + k'_0 P^k x^j} + (S_1 - S_0) \frac{k_1{}^n (P_k X_j) x^n}{(P_k X_j) + k_1{}^n (P_k X_j) x^n} \quad (9)$$

Collection of terms yields:

$$S_{app} = \frac{S_0}{1 + k_1{}^n x^n} + \frac{S_1 k_1{}^n x^n}{1 + k_1{}^n x^n} \quad (10)$$

It can be seen that Eq. (10) is now in the same form as Eq. (4) and that the association parameter k'_0 has canceled out. Thus, only binding sites linked to changes in solubility will be seen in this analysis. This result is in line with the experimental data collected by Waugh et al. (1971), which also showed through sedimentation analysis that an associated aggregate and not the monomer participates in the precipitation reaction.

Salt-induced solubility profiles were directly analyzed using a Gauss-Newton nonlinear regression analysis program developed at this laboratory by Dr. William Damert. All profiles were analyzed by fixing the values of n and m and calculating the best least squares fit for the optimum evaluated k_1 and k_2 values. The n and m values were then fixed to new integer values and the entire procedure was repeated. The n and m values that yielded the minimum root-mean-square value for the analysis with the minimum error in k_1 and k_2 were then reported.

Molecular Modeling

All aggregate structures employed the α_s-casein B casein monomer structure previously refined via energy minimization (Kumosinski et al., 1994). Aggregates were constructed using a docking procedure on an Evans and Sutherland PS390 interactive computer graphics display driven by the Tripos Sybyl (St. Louis, MO) molecular modeling software on a Silicon Graphics 4200 Unix-based computer. The docking procedure allowed for individual manipulation of the orientation of up to four molecular entities relative to one another. The desired orientations could then be frozen in space and merged into one entity for further energy minimization calculations utilizing a molecular force field. The criterion for acceptance of reasonable structures was determined by a combination of experimentally determined information and the calculation of the lowest energy for that structure.

Force Field Calculation. Studies concerned with the structures and/or energetics of molecules at the atomic level require a detailed knowledge of the potential energy surface (i.e., the potential energy as a function

of the atomic coordinates). For systems with a small number of atoms, quantum mechanical methods may be used, but these methods become computationally intractable for larger systems (e.g., most systems of biological interest) because of the large number of atoms that must be considered. For these larger systems, molecular mechanics methods are used. Molecular mechanics is based on the assumption that the true potential energy surface can be approximated with an empirical potential surface consisting of simple analytical functions of the atomic coordinates. The empirical potential energy model treats the atoms as a collection of point masses that are coupled to one another through covalent (bonded) and noncovalent (nonbonded) interactions. The potential energy function (Weiner et al., 1986; Kollman, 1987) generally has the form:

$$\begin{aligned} E_{\text{total}} = & \sum_{bonds} K_r(r - r_{eq})^2 + \sum_{angles} K_\theta(\theta - \theta_{eq})^2 \\ & + \sum_{dihedrals} \frac{1}{2}K\,[1 + \cos(n\phi - \gamma)] \\ & + \sum_{i<j} \frac{B_{ij}}{R_{ij}^{12}} - \frac{A_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\epsilon R_{ij}} \end{aligned} \tag{11}$$

The first three terms are due to covalent interactions and represent the difference in energy between the geometry of the actual structure and a geometry in which the bond lengths, bond angles, and dihedral angles all have ideal values. The remaining terms represent nonbonded van der Waals and electrostatic interactions. In Eq. (11), r, θ, ϕ, and R_{ij} are variables, determined by the atomic coordinates. All other entities are constant parameters chosen to reproduce experimental observables as closely as possible. Although empirical potential energy functions such as Eq. (11) are relatively crude, they have been applied successfully to the study of hydrocarbons, oligonucleotides, peptides, and amino acids, as well as systems containing a large number of small molecules such as water. The Tripos force field in Tripos' Sybyl software package uses the above functional form, plus a bump factor, which allows atoms within a fraction of the van der Waals radius for H-bond function if so chosen by the user. The parameters used for electrostatic calculations include atomic partial charges (q_i) calculated by the Kollman group (Weiner et al., 1986; Kollman, 1987) using a united atom approach with only essential hydrogens. All molecular structures were refined with an energy-minimization procedure using a conjugate gradient algorithm, in which the positions of the atoms are adjusted iteratively so as to achieve a minimum potential energy value. Energy-minimization calculations were termi-

nated when the energy difference between the current and previous iterations was less than 1 kcal/mol. A nonbonded cutoff of 5 Å was used initially to save computer time and is an appropriate value for use of a function that varies with distance. A stabilization energy of at least -10 kcal/mol/residue was achieved for all structures, which is consistent with values obtained for energy minimized structures determined by x-ray crystallography.

Molecular Dynamics. In the previous paragraph we considered only static structures. However, the dynamic motion of molecules in solution contributes to their functionality. The molecular dynamics approach is a method of studying motion and molecular configuration as a function of time (Andersen, 1980). All atoms in the molecule are assigned a kinetic energy through a velocity term, which can be related to the local temperature as well as to the average temperature of the system. These calculations can be performed in vacuum or in the presence of a desired number of solvent molecules such as water and at a constant temperature and volume using a periodic boundary condition to confine the calculation within a prescribed volume (van Gunsteren and Berendsen, 1977). For these calculations a force field describing the potential energy is combined with Newton's second law of motion.

$$F_i = m_i a_i(t) = m_i \frac{dv_i(t)}{dt} = m_i \frac{d^2x_i(t)}{dt^2} = -\nabla_i E \tag{12}$$

where F_i is the force on atom i, which has mass (m_i), velocity (v_i), acceleration (a_i), and position (x_i). ∇_i is the gradient or the derivative with respect to position, t is the time displacement, and E is the potential energy of the molecule described by the chosen force field. Equation (12) is integrated at various time intervals for the desired molecule using the chosen force field via a prescribed numerical integration method. The time interval chosen must be small in comparison with the period associated with highest frequency of motion within the molecule. This is usually stretching of a bond associated with a hydrogen atom, i.e., one femtosecond. Numerical integration of Eq. (12) over 1-fsec intervals to 100 psec for a protein molecule of 2000 atoms or more necessitates a fast computer with a large memory capacity. The results of these calculations can mimic the motions of molecules in solution and also time-dependent geometric parameters. For example, the distance from the center of moment for a set of atoms may be related to correlation times derived from NMR, EPR, and fluorescence experiments.

QUANTITATION OF THE SALT-INDUCED SOLUBILITY PROFILES OF THE CASEINS

Solubility at 37°C

Solubility determinations of α_{s1}-caseins A and B (α_{s1}-A, α_{s1}-B) and β-casein C (β-C) were performed at 37°C in 10 mM imidazole-HCl pH 7.0, 0.07 M KCl, at initial protein concentrations of 10 mg/mL (Farrell et al., 1988). As in the experiments of Noble and Waugh (1965), the proteins precipitate when added $CaCl_2$ exceeds 5 mM (Fig. 2.1A and B). Creamer and Waugh (1966) had suggested that about 13 sites of similar calcium binding strength exist in the α_{s1}-B at pH 6.6 and that when calcium ion concentration exceeds this critical binding level, charge neutralization occurs and precipitation results. Comparison of the solubility profiles of α_{s1}-A and -B indicates that at 37°C α_{s1}-A is more soluble than α_{s1}-B, while β-C is the most soluble. In order to quantify the data, nonlinear regression analyses were performed. The data of Fig. 2.1A were fitted by Eq. (4). Values of k_1 were obtained at fixed integer values of n; the correct value of n was taken to be the fit with the minimum root mean square (RMS). Fig. 2.1A shows the fit to $n = 2$, 4, and 8 for α_{s1}-A; values for $n = 8$ gave the minimum RMS with the lowest error in k_1. Analysis of the solubility profiles of α_{s1}-A, α_{s1}-B, and β-C at 37°C, where hydrophobic interactions are maximized, showed no salting-in behavior so that k_2 and m were essentially zero. Values obtained for k_1 (salting-out) and n are given in Table 2.1.

Solubility at 1°C

Figure 2.2 shows that β-casein C is not precipitated at 1°C by Ca^{2+} at concentrations of up to 400 mM. It is known that hydrophobic forces are dominant in the association reactions of β-caseins (Schmidt, 1982). The solubility of β-casein C clearly distinguishes it from α_{s1}-B; it is known that β-casein binds Ca^{2+} at 1°C, but in this case binding is not linked to changes in solubility and so no analysis by thermodynamic linkage is possible. Aliquot addition of calcium chloride solutions to α_{s1}-casein results in a rapid decrease in solubility from 8 to 50 mM, where the protein is almost totally precipitated. When the calcium chloride concentration exceeds 100 mM, a gradual salting-in of the protein ensues at 1°C. The data for α_{s1}-B were fitted by Eq. (4) and the salting-out parameters k_1 and n, as well as the salting-in parameters k_2 and m were determined (Table 2.2). The α_{s1}-A, genetic variant, in contrast to the α_{s1}-B exhibits extraordinary solubility behavior over a broad range of calcium chloride concentrations. At 1°C (Fig. 2.3) α_{s1}-A, like α_{s1}-B (Fig. 2.2), is precipitated

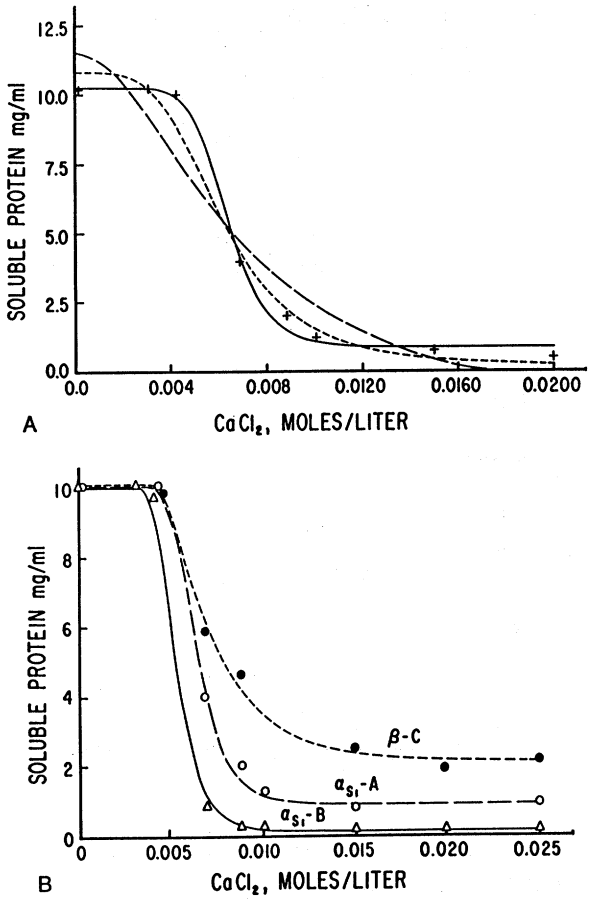


FIG. 2.1 Solubility at 37°C of the calcium salts of α_1-caseins A and B and β-casein C as a function of increasing $CaCl_2$ concentration. Solutions buffered at pH 7.0, 10 mM imidazole-HCl. (A) The experimental data for α_{s1}-A were fitted by Eq. (4) by nonlinear regression analysis with values of 2 (— —), 4 (---), and 8 (—) assigned to n. The best fit was obtained for $n = 8$. (B) Similar fits for α_{s1}-B and β-casein C; results of analyses are shown in Table 2.1.

with calcium at about 8 mM, whereupon the net electrical charge on the Ca-complexed protein may be close to zero. In the absence of electrolyte (KCl) or buffer, and after aliquot addition of $CaCl_2$, the protein is driven into solution at 90 mM. The Ca-complexed protein is now positively

TABLE 2.1 Calcium-Induced Insolubility of Casein at 37°C[a]

Casein	k_1(L/mole)	n	S_1[b](mg/mL)
α_{s1}-A	157 ± 3	8	0.9 ± 0.2
α_{s1}-B	186 ± 3	8	0.1 ± 0.1
β-C	156 ± 12	4	2.0 ± 0.3

[a]Solutions buffered at pH 7.0, 10 mM imidazole-HCL, 0.07M KCl.
[b]S_1 denotes the maximum value for soluble protein at elevated Ca^{2+} concentrations.

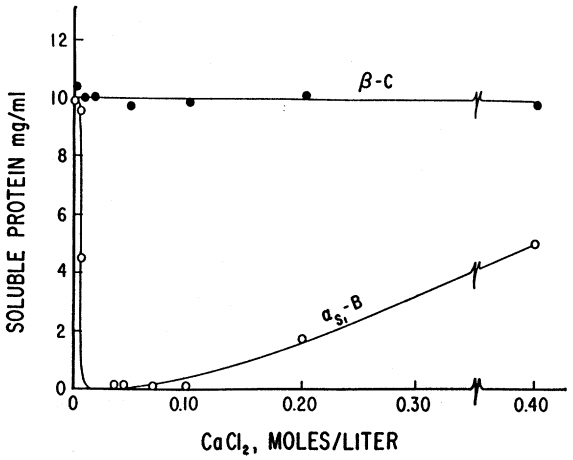


FIG. 2.2 Solubility at 1°C of calcium α_{s1}-B caseinate and calcium β-C caseinate as a function of increasing $CaCl_2$ concentration. Data were fitted by Eq. (4); results of analyses are shown in Table 2.2.

charged and is acting as a cation. This conclusion was verified by free-boundary electrophoresis at pH 7.0, 10 mM imidazole, 150 mM $CaCl_2$, where the protein is soluble at 1°C; it migrates (+ 1.36 cm^2 $volt^{-1}$ sec^{-1} × 10^{-5}) toward the cathode (Thompson et al., 1969). The evidence to this point favors direct salt-protein interactions as being responsible for the salting-in and salting-out of caseins by calcium rather than a salt-solvent interaction as previously proposed by Melander and Horvath (1977) following the cavity theory model of Sinanoglu (1968). The concept of salt binding is supported by the earlier work of Robinson and Jencks (1965), who studied salting-out of model compounds and concluded that binding was a factor.

TABLE 2.2 Calcium-Induced Insolubility and Solubility of Caseins at 1°C[a]

Casein	k_1[b]	n	k_2[b]	m
α_{s1}-B	123 ± 5	8	2.5 ± 0.2	4
α_{s1}-A	68 ± 1	8	10.6 ± 0.3	8
β-C	Totally soluble			

[a]Conditions as in Table 2.1.
[b]L/mole.

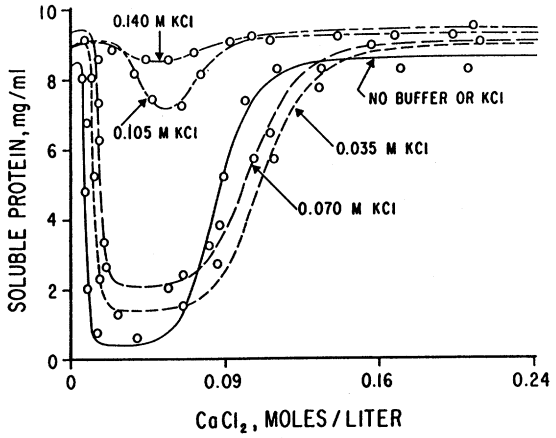


FIG. 2.3 Solubility at 1°C of calcium α_{s1}-A caseinate as a function of increasing $CaCl_2$ and KCl concentrations. Data were fitted by Eq. (4); results are shown in Table 2.3.

Influence of Electrolyte on Salting-Out and Salting-In Constants

Aliquot addition of $CaCl_2$ in the presence of various concentrations of KCl results in a shift in the solubility profile of α_{s1}-A at 1°C (Fig. 2.3). KCl was chosen as the electrolyte because it occurs at higher concentration than NaCl in milk serum (Farrell and Thompson, 1987). The data were analyzed and fitted by Eq. (4); parameters k_1, k_2, n, and m are given in Table 2.3. Note that as the KCl concentration increases, n drops to 4 for Ca^{2+}; the protein apparently requires less bound Ca^{2+} to precipitate. On the other hand, m for resolubilization remains at 8. Increasing ionic

TABLE 2.3 Ionic Strength Dependence of Calcium-Induced Solubility of α_{s1} – casein A at 1°C

KCl mM	k_1[a]	k_2[a]	S_1[b](mg/mL)
—	130 ± 3	13.3 ± 0.3	0.4 ± 0.2
35[c]	82 ± 1	10.1 ± 0.1	1.4 ± 0.1
70[c]	68 ± 1	10.6 ± 0.2	2.1 ± 0.1
105[d]	22 ± 1	15.9 ± 0.1	4.7 ± 0.4
140[d]	34 ± 1	15.7 ± 0.2	8.4 ± 0.1

[a]L/mole.
[b]S_1 denotes maximum value for soluble protein after precipitation but before total resolubilization.
[c]$n = 8$ and $m = 8$.
[d]$n = 4$ and $m = 8$.

strength by the addition of KCl for each calcium-induced solubility profile of α_{s1}-A at 1°C decreased k_1 and k_2 values. The variations of k_1 and k_2 with KCl are given in Fig. 2.4A and B. For k_1, a nearly monotonic decrease occurs; Fig. 2.4A can be analyzed as a potassium-binding isotherm itself and an association constant K_{a1} computed for KCl-protein interactions. The value given in Table 2.4 line 1 was found to be 20 ± 6 L/mole. For k_2, the variation appeared more complex with two transitions. Analyzing these data, K_{a2} and K_{a2}^1 were calculated (Table 2.4, lines 2 and 3). As compared with the primary k_1 values for Ca^{2+} precipitation, all derived constants related to KCl effects are substantially smaller. Changes are consistent with the electrostatic character of these binding isotherms and could be attributed to a competitive effect of elevated potassium ion for calcium-binding sites or to KCl-solvent interactions affecting solubility; again for this system direct salt binding to the caseins appears to be the driving force. The overall difference for the behavior of α_{s1}-A relative to α_{s1}-B, however, must reside in the structural differences due to the deletion mutation in the A variant. In all of the above analysis, the constants obtained approximate binding constants, but they may not be identical with average association constants, since only binding that leads to changes in solubility is disclosed by this analysis. Bear in mind that for β-casein (Fig. 2.2) binding occurs but no change in solubility accompanies this and so no analysis is possible.

Influence of Electrolyte on Soluble Protein (S_1)

Further insights into the source of the effects resulting in the α_{s1}-A solubility profiles can be obtained by analysis of the variation of the

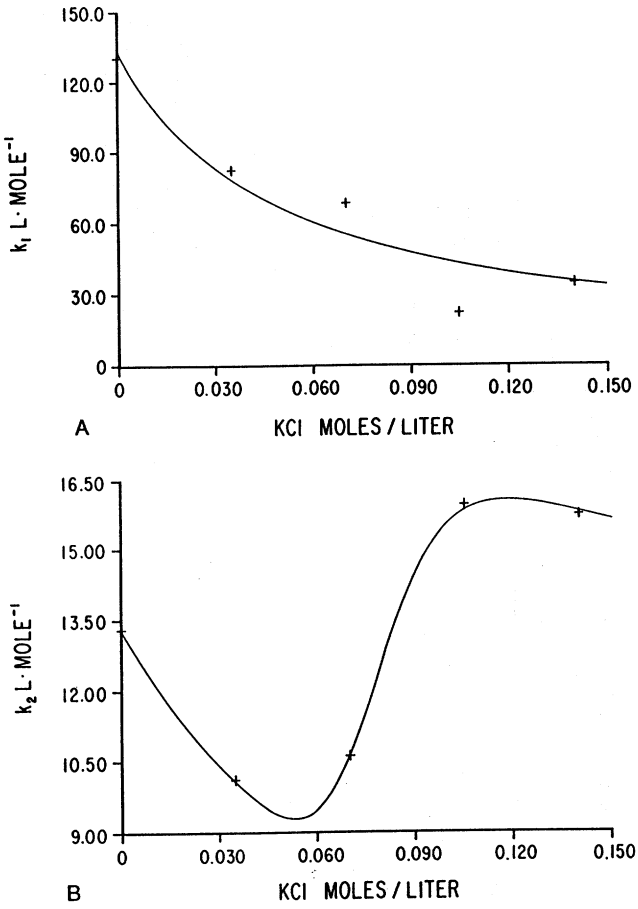


FIG. 2.4 Variations of k_1 (A) and k_2 (B) with KCl concentration. These parameters obtained from analysis of Fig. 2.3 (Table 2.3) were fitted by Eq. (4); results are shown in Table 2.4.

parameter S_1 of Table 2.3. This term is taken to represent the soluble form of α_{s1}-A in the valley between 2 and 8 mM $CaCl_2$ (Fig. 2.3).

Since the magnitude of salting-in for S_1 is larger than expected for weaker salt-solvent forces, it was decided to analyze the variations of S_1 with KCl as a binding isotherm (Fig. 2.5). Two cooperative transitions appear to occur. These were analyzed in terms of Eq. (4) and values of K_{as}, K_{as}^1 and n were calculated; these derived constants are compared with those derived from the variance of k_1 and k_2 with KCl in Table 2.4. All of the constants derived from k_2 or S_1 appear to be self-consistent in

TABLE 2.4 Comparison of Derived Constant from the Effect of KCl on k_1, k_2, and S_1 from the Calcium-Induced Solubility of α_{s1}-A at 1°C

Derived constant	Source of constant	Value (L/mole)	n	$\log K$
K_{a1}	k_1[a]	20.0 ± 6.0	1	1.30
K_{a2}	k_2-1st transition[b]	9.2 ± 0.6	1	0.96
K_{a2}^1	k_2-2nd transition[b]	12.5 ± 0.5	8	1.10
K_{as}	S_1-1st transition[c]	15.9 ± 0.8	1	1.20
K_{as}^1	S_1-2nd transition[c]	8.9 ± 0.3	8	0.94
K_a	$K^+ + HPO_4^{-2}$[d]	6.3	1	0.80

[a]See graph Fig. 2.4A.
[b]See graph Fig. 2.4B.
[c]See graph Fig. 2.5.
[d]Sillen and Martell (1971).

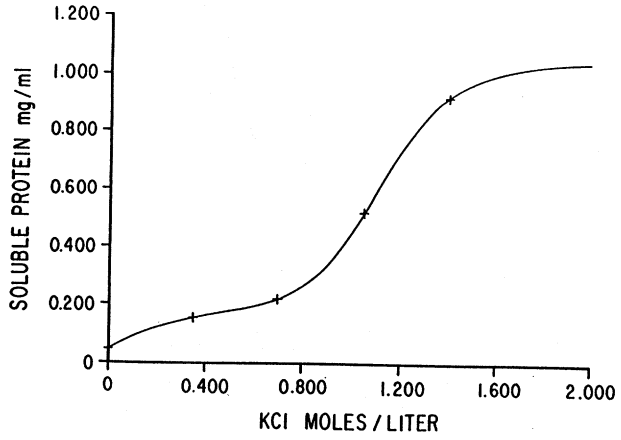


FIG. 2.5 Effect of KCl concentration (M × 10) on change of solubility of protein (α_{s1}-A)at 1°C relative to solubility in the absence of KCl. Data from Table 2.3 were fitted by Eq. (4); results are shown in Table 2.4.

order of magnitude and in effect (salting-in) and are in the order of magnitude of K_a for $K^+ + HPO_4^{-2}$. The most readily apparent conclusion is that salting-in, in this particular case, may be the result of direct salt (perhaps both anion and cation) interactions with charged groups (possibly phosphates) of the protein. This effect, as noted above, has

somewhat wide-ranging implications. It had previously been postulated (Horvath and Mellander, 1977) that such calcium-protein-salt interactions occur primarily through salt-solvent interactions rather than direct salt binding to protein.

Effect of Other Cations at 1°C

Figure 2.6 illustrates the solubility of α_{s1}-A in the presence of various cations. Cu^{2+} and Zn^{2+} are the most effective precipitants. Coordinate complexes may be formed between α_{s1}-A molecules with Cu^{2+} and Zn^{2+} and Co^{2+}. Ca^{2+} is effective as a precipitant to a lesser extent than Cu^{2+} or Zn^{2+}, whereas Mg^{2+} is the least effective of the five cations studied. The salting-out and salting-in constants were estimated for each cation and are given in Table 2.5. Cation variation (i.e., use of magnesium, calcium, cobalt, copper, and zinc) of these profiles showed k_1 and k_2 behavior consistent with concepts of phosphate- and carboxylate-ligand coordination, respectively. Clearly, an inverse relationship exists between casein solubility (as quantified by changes in k_1) and the atomic number of the divalent cations studied. The salting-in constant k_2 appears to decrease and then increase with atomic number; no apparent correlation with ionic radius is evident, so that the best possible explanation is that as the atomic weight of the cation increases, the solubility decreases.

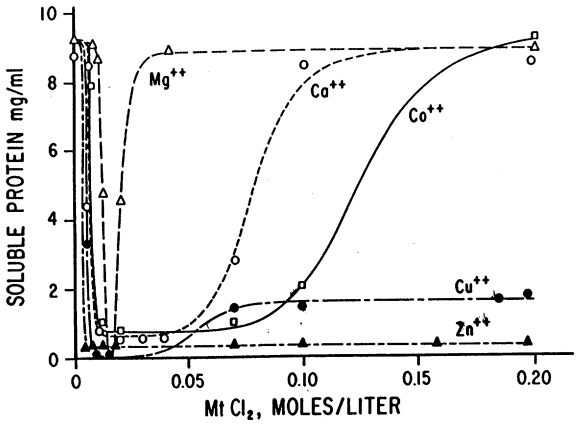


FIG. 2.6 Solubility at 1°C of various salts of α_{s1}-A caseinates as function of increasing concentration. Data were fitted by Eq. (4); results are given in Table 2.5.

TABLE 2.5 Cation-Induced Solubility of α_{S1}-Casein A at 1°C[a]

Cation	k_1[b]	k_2[b]	$(S_2 - S_1)$[c] (mg/mL)	Atomic no.	R[d](Å)
Mg^{2+}	76 ± 9	56 ± 7	8.4 ± 1.3	12	0.66
Ca^{2+}	150 ± 27	13 ± 2	8.2 ± 1.4	20	0.99
Co^{2+}	166 ± 4	8.1 ± 0.2	8.5 ± 0.2	27	0.72
Cu^{2+}	229 ± 2	18 ± 2	1.5 ± 0.1	29	0.72
Zn^{2+}	373 ± 27	202 ±35	0.20 ± 0.03	30	0.74

[a] $n = m = 8$ for all calculations.
[b] L/mole.
[c] $(S_2 - S_1)$ = concentration of soluble α_{s1}-A in mg/mL.
[d] Cation atomic radius in Å.

Influence of Phosphate Groups on Salting-Out and Salting-In

As shown in Fig. 2.1, α_{s1}-A and -B and β-C readily precipitate at 37°C in 0.07 M KCl. Since under these conditions n was correlated with the number of phosphate residues in the native casein, the importance of these residues in the precipitation reaction could be tested. In previously conducted research, the phosphate groups of α_{s1}-B were removed enzymatically (Bingham et al., 1972) and the effects of KCl on the precipitation of native (N) and dephosphorylated (O-P) caseins had been compared but not quantitated (Fig. 2.7A shows a typical curve from Bingham et al., 1972). Analysis of these data by use of Eq. (4) (Farrell and Kumosinski, 1988) is summarized in Table 2.6. With no KCl present, dephosphorylation increases k_1 and some salting-in occurs for the O-P form; surprisingly, for both proteins (N and O-P) $n = 16$. Also, in the absence of KCl, salting-in occurs only for the dephosphorylated α_{s1}-casein B.

When α_{s1}-A is dephosphorylated, it becomes nearly completely soluble at 1°C and is salted-in even at 36°C (Fig. 2.7B). In contrast (Fig. 2.1), the native α_{s1}-A is not appreciably salted-in at 37°C. Results are compared in Table 2.6. For dephosphorylated α_{s1}-A without added KCl $n = 8$, mirroring the numbers found for the native protein, but for the O-P form at 1°C $n = 2$ and $m = 4$. The k_2 values observed for both native and O-P α_{s1}-A at both 1 and 36°C are similar to each other and to that of α_{s1}-B at 1°C. However, the small degree of salting-in that occurs for O-P of α_{s1}-B does so with an elevated k_2 (Table 2.6), showing another significant difference between the A and B variants.

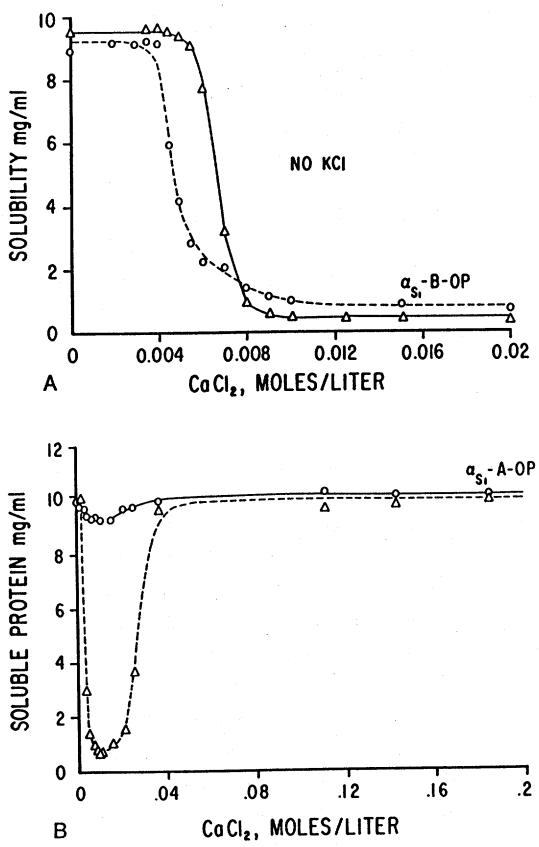


FIG. 2.7 (A) Solubility at 37°C of native and dephosphorylated α_{s1}-B (α_{s1}-B–OP) as a function of $CaCl_2$ concentrations at 10 mg/mL. (B) Solubility of α_{s1}-A O-P as a function of calcium ion concentrations at 1°C (—) and at 36°C (- - -). Data were fitted by Eq. (4). Results of analyses are given in Table 2.6.

Calcium-Induced Protein Self-Association Model

Finally, it would be advantageous to test the salt-induced protein self-association to determine whether or not the monomer of α_{s1}-casein is the molecular unit responsible for the salting-out process. Here, we used the sedimentation data of Waugh et al. (1971) at low calcium chloride concentrations at 37°C and at a protein concentration of 10 mg/mL.

TABLE 2.6 Calcium-Induced Solubility of Native (N) and Dephosphorylated (O-P) α_{s1}-Casein B (B) and α_{s1}-Casein A (A)

Protein	Temp (°C)	k_1 (L/mole)	k_2 (L/mole)	S_1	S_2	n	m
NB	37	151 ± 1		0.49 ± 0.04		16	
NBO-P	37	219 ± 2	135 ± 12	2.6 ± 0.3	0.8 ± 0.06	16	8
NA	36	140 ± 3		0.9 ± 0.2		8	
NAO-P	36	326 ± 7	36 ± 1	0.6 ± 0.3	10.0 ± 0.6	8	8
NA	1	130 ± 4	13 ± 1	0.4 ± 0.2		8	
NAO-P	1	223 ± 59	46 ± 4	8.8 ± 1.8	10.0 ± 0.6	2	4

Two distinct peaks were observed, and Fig. 2.8 shows the variation of sedimentation coefficient $S_{20,w}$ for the slow peak (Fig. 2.8A), and fast peak (Fig. 2.8B) with added $CaCl_2$ up to a salt concentration of 0.007 M. At greater concentrations the protein precipitates. The appearance of a bimodal pattern in the sedimentation profile as the $CaCl_2$ is increased is indicative of a ligand-induced protein self-association system under equilibrium conditions (Cann, 1978). Thermodynamic linkage and nonlinear regression analysis of the data using Eq. (4), but substituting the sedimentation constant(s) for S_{app}, was carried out. For the slow peak, Fig. 2.8A, a biphasic binding process was determined, whereas analysis of the fast peak, Fig. 2.8B, yields only one cooperative binding mechanism leading to protein self-association. The results of this analysis of Fig. 2.8 are given in Table 2.7. Here, for the slow peak the k_1 and k_2 values are 300 ± 50 and 166 ± 34 L/mole, respectively, with $n = 1$ and $m = 8$. The value of S_0 of 1.7 ± 0.3 for the slow peak is of the correct order of magnitude for α_{s1}-casein monomer with molecular weight of 24,000. The S_1 and S_2 values were 4.6 ± 0.5 and 12.6 ± 1.1 S, respectively. If we assume that S_0 is the sedimentation coefficient of monomeric α_{s1}-casein B, the degree of self-association can be approximated by dividing the S_1 and S_2 values by S_0 and raising that quotient to the 1.5 power, which yields values of 4.3 and 18.8 for the size of the calcium-induced aggregates of the slow peak (Cann, 1978). Using the same algorithm, an approximate degree of self-association of 8.18 and 46.5 can be calculated from S_0 and S_1 values of 7.4 ± 0.5 and 22.5 ± 1.5 S, respectively, obtained from thermodynamic linkage analysis of the fast component (Table 2.7). Thus, from the slow peak the apparent binding of one mole of $CaCl_2$ to α_{s1}-casein induces tetramer formation, which in turn leads to 18-mer when 8 moles of $CaCl_2$ bind cooperatively. (*Note*: Thermody-

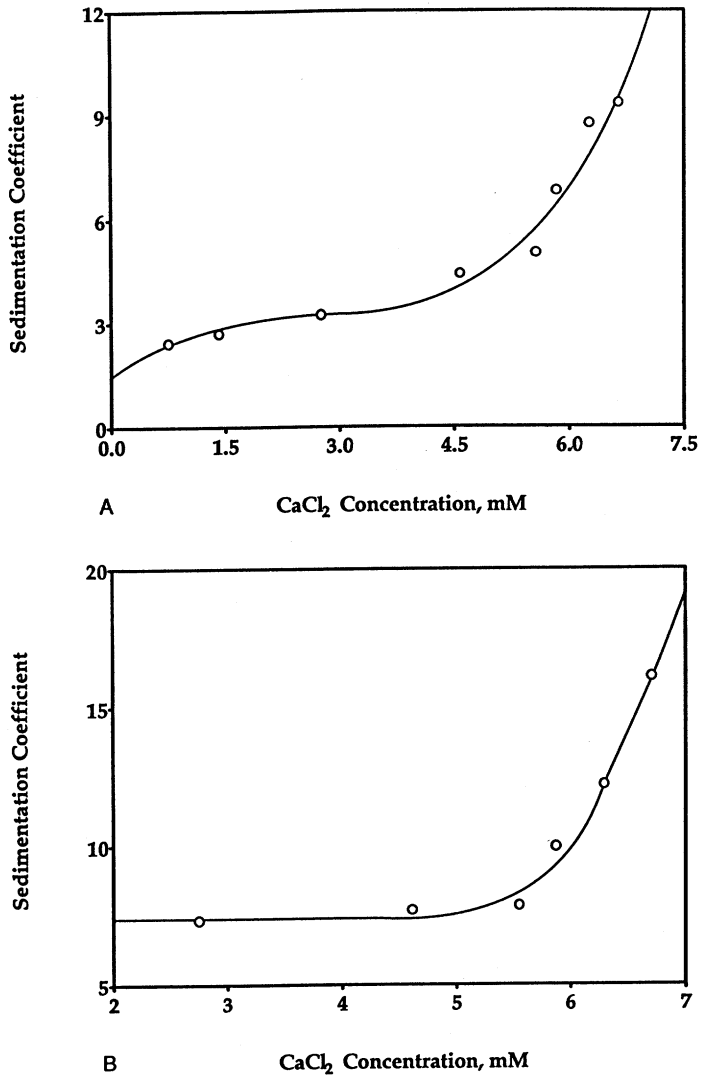


FIG. 2.8 Sedimentation velocity of α_{s1}-B at 10 mg/mL as a function of added $CaCl_2$. Two peaks were observed: a slow one (A) and a fast one (B). Analysis was by use of Eq. (4) and is given in Table 2.7.

TABLE 2.7 Thermodynamic Linkage Results of Sedimentation Coefficient of α_{s1}-Casein B vs $CaCl_2$[a]

Peaks	k_1(L/mole)	k_2(L/mole)	S_0	S_1	S_2
Slow[b]	300	166	1.7	4.6	12.6
	±50	±34	±.3	±.5	±1.1
Fast[c]	152		7.4	22.5	
	±32		±.5	±1.5	

[a]At 37°C with 10 mg/mL of protein.
[b]$n = 1$, $m = 8$.
[c]$n = 16$.

namic linkage reveals only those binding events that correlate with the observed physical changes.) The fast peak with a S_0 value of 7.4 or an octamer then appears to bind 16 moles of $CaCl_2$ cooperatively to yield a 46-mer aggregate with an S value of 22.5 ± 1.5. The analysis clearly demonstrates that α_{s1}-casein B becomes insoluble as an aggregate and not as a monomer unit. It is also interesting to observe that the k_2 value of the slow peak and the k_1 of the fast peak (Table 2.7) are 166 and 152 L/mol, respectively, and these values are comparable to the salting-out binding parameters ($k_1 = 186$) seen in Table 2.1. In contrast k_1 of the slow peak approaches the average K_A for binding of Ca^{2+} to α_{s1}-casein (380 L/mole) (Dickson and Perkins, 1971).

MOLECULAR MODELING OF α_{s1}-CASEIN

In the previous sections we have formed a quantitative thermodynamic mechanism for the salt-induced solubility profiles of α_{s1}-casein as a function of temperature, genetic variation, dephosphorylation, i.e., post-translational modification, and types of divalent cations employed. Here, the insolubility portion was described as a cooperative binding of n moles of salt with a constant of k_1 leading from a protein solubility of S_0 at zero salt concentration to a value of S_1. The salting-in process was quantitatively described by m moles of salt cooperatively binding to the protein with a constant of k_2 and yielding a limiting protein solubility of S_2. However, it must be stressed that the ligand-induced protein association model proposed by Cann and Hinmen (1976) cannot be used to describe the salt-induced precipitation profile used in this study (see theory section). Only molecular weight or sedimentation coefficient studies at lower salt concentrations such as those performed by Waugh et al. (1971) can

totally describe the salt-induced self-association model as shown in the previous section. Hence, it would be prudent to assume that the oligomeric unit of α_{s1}-casein that leads to precipitation is a tetramer or octamer as concluded above.

In the molecular modeling section, we now attempt to construct tetramer and octamer models from an energy-minimized, predicted, three-dimensional structure of α_{s1}-casein (Kumosinski et al., 1994). Here, the antiparallel stranded sheets that occur in the hydrophobic portion of α_{s1}-B will be used as interaction sites for the construction of these tetrameric and octameric models. It is hoped that the presentation of such predicted structures will aid researchers in designing new chemical and biochemical modification experiments to test these working models.

To study the salt-induced solubility profiles of α_{s1}-casein by modeling, the A variant was chosen and its energy-minimized molecular model produced by deletion of the appropriate residues from the α_{s1}-casein B model. The α_{s1}-casein A structure was chosen to describe the salt-induced salting-in profiles since, as shown in the previous section, this genetic variant of α_{s1}-casein would be salted-in with added $CaCl_2$ at 1°C. Also, since molecular dynamic (MD) calculations in the presence of salt and water are very time consuming, it is more efficient to utilize only that portion of the molecule to which salt would most probably bind, i.e., the hydrophilic domain. Therefore, only residues 1 to 86 of α_{s1}-casein A were used in these simulations. In addition, MD calculations were performed on a dephosphorylated molecule to mimic the effect of phosphoserines on salt-induced solubility profiles.

Hydrophobic and Hydrophilic Dimers and Oligomers

The energy-minimized structure generated for α_{s1}-casein B as described above is shown in Fig. 2.9A, where it is displayed from carboxy- to amino-terminal (left to right). Analysis of this structure shows the molecule to be composed (right to left) of a short hydrophilic amino-terminal portion, a segment of rather hydrophobic β-sheet, the phosphopeptide region, and a short portion of α-helix, which connects this N-terminal portion to the very hydrophobic carboxyl-terminal domain containing extended antiparallel β-strands. For clarity the backbone without side chains is shown in Fig. 2.9A with prolines (P) indicated, and an accompanying chain trace stereo view (Fig. 2.9B) is given.

From the overall shape of the α_{s1}-B model (Fig. 2.9), it is apparent that it is impossible to approximate its structure with either a prolate or an oblate ellipsoid of revolution, as was done in the case of the β-casein refined structure (Kumosinski et al., 1993). Indeed, a rather large degree

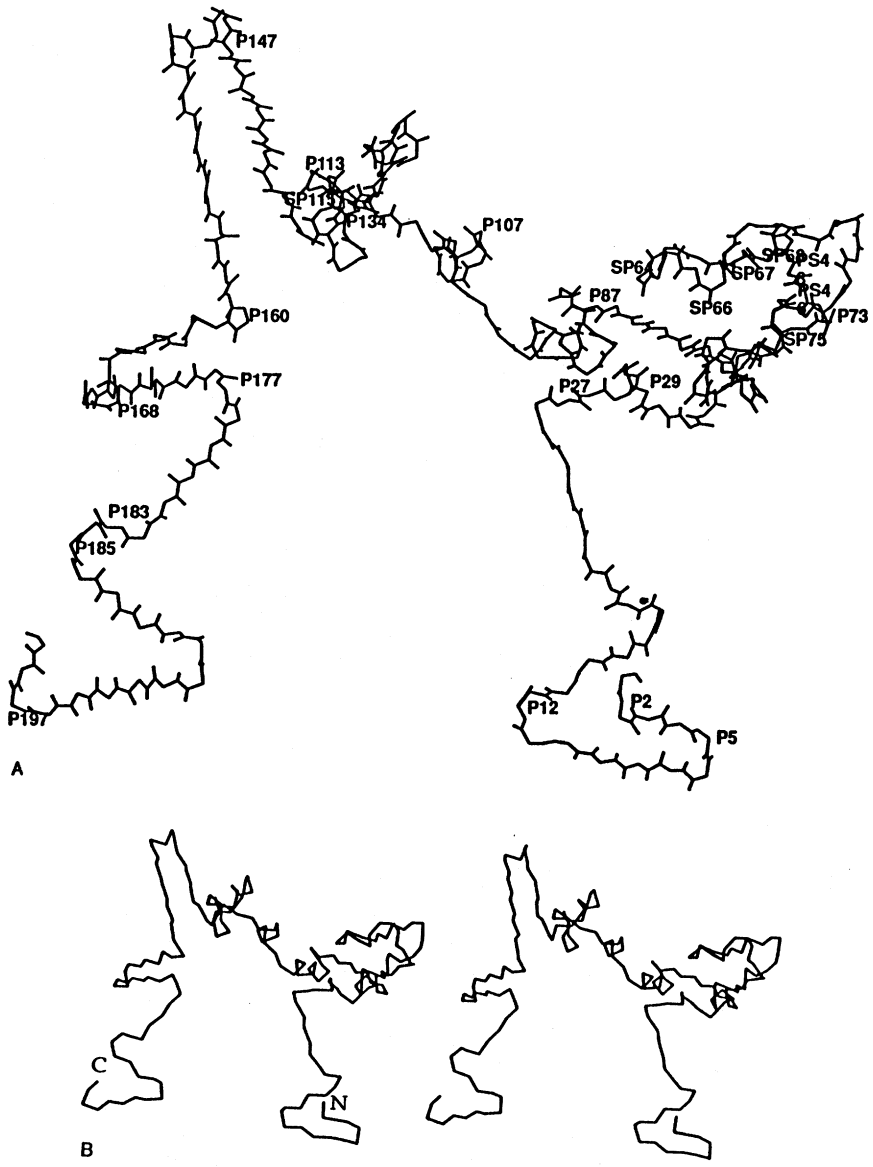


FIG. 2.9 (A) Backbone of the refined model of α_{s1}-casein, without side chains; prolines (P) and serine phosphates (SP) indicated. (B) Stereo view of the refined three-dimensional molecular model of α_{s1}-casein; the N- and C-terminal ends of the molecule are labeled.

of asymmetry is observed. As noted above, the hydrophilic and hydrophobic domains are joined by extended structures, whose central feature is an α-helix with its pitch perpendicular to the two domains. It is speculated that this α-helix would be important for preserving the integrity of the two domains when a dynamic calculation is finally performed.

Schmidt (1982) has summarized the light-scattering studies on variants of α_{s1}-casein under a variety of environmental conditions from which a stoichiometry of the α_{s1}-casein self-association is obtained at selected temperatures and ionic strengths. From the results, it can be concluded that α_{s1}-casein undergoes a concentration-dependent, reversible, hydrophobically controlled association from monomer to dimer then tetramer, hexamer, octamer, and even higher if the ionic strength is increased, especially as shown (Fig. 2.8) in the case of added $CaCl_2$. To mimic hydrophobic salt-induced self-association mechanisms, we now attempt to construct an energy-minimized dimer, tetramer, and octamer structure using hydrophobic and hydrophilic sites.

The first step is to create a dimer from the large-stranded β-sheets, which occur at residues 136–158. The side chains are predominately hydrophobic, and the hydrogen bonding of the sheet secondary structure yields rigidity to this site. After docking, a dimer can easily be formed if two of these stranded sheets are docked in an antiparallel fashion (Fig. 2.10A). Such an asymmetric arrangement minimizes the dipole-dipole interactions of the backbones while allowing the hydrophobic side chains to interact freely. In fact, this structure, following minimization (Fig. 2.10A), displays a stabilizing energy of -452 kcal/mole/residue, over monomer, i.e., $-520 = E_2 - 2 \cdot E_m$ where E_2 is the energy of the dimer and E_m is the monomer energy.

Another possible interaction site for hydrophobic dimerization resides is the deletion peptide of α_{s1}-casein A (i.e., the peptide that is deleted from α_{s1}-casein B to form α_{s1}-casein A). Closer inspection of the residues of this peptide show a β-sheet secondary structure with hydrophobic as well as acidic and basic side chains. Thus, by docking two molecules in an antiparallel fashion, a hydrophobically stabilized intermolecular ion pair (Tanford, 1967) can be formed upon the construction of a dimer. The dimer was then energy minimized and is shown in Fig. 2.10B. The formation of this hydrophobic ion pair has been used to explain the differences in the calcium-induced solubility and colloidal stability between α_{s1}-caseins B and A (Farrell et al., 1988; Kumosinski and Farrell, 1991). Small-angle x-ray scattering of micelles reconstituted from whole caseins containing α_{s1}-B or -A shows large differences with respect to

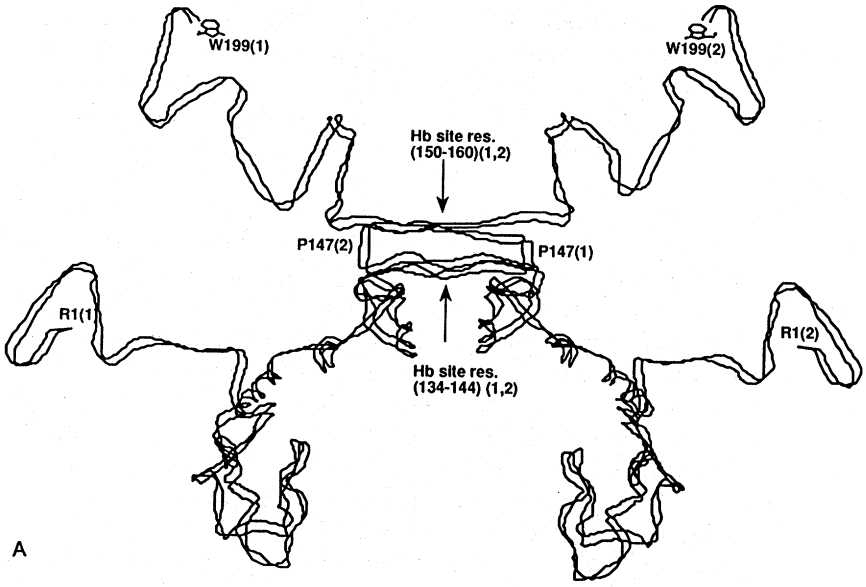
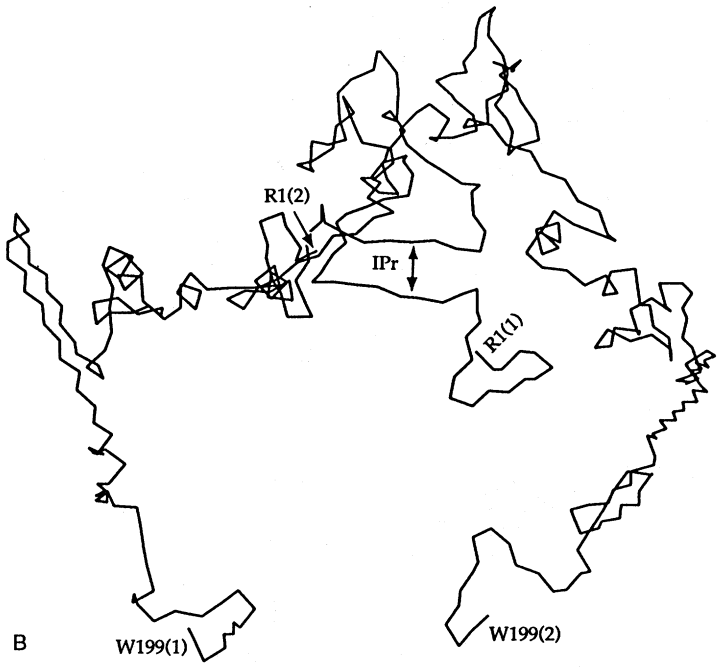


FIG. 2.10 (A) α-Carbon chain trace of backbone without side chains of hydrophobic stabilized (Hb) antiparallel sheet dimer; the large sheets centering on proline 147 are docked. (B) Backbone structure of hydrophobic ion-pair dimer (IPr) from α_{s1}-casein B; this area contains the α_{s1}-A deletion peptide. Key for labels: $R_1(1)$, W199 (1) represent N and C terminals of molecule 1 (arginine 1 and tryptophan 199 of molecule 1; the 2 in parentheses refers to molecule 2. Both dimeric structures energy minimized to -10 kcal/mole/residue.

submicellar packing density within the micelle structure, i.e., 3 to 1 versus 6 to 1 for B and A, respectively. This difference in packing density may be a result of destructive interference in the scattered intensity due to an asymmetrical structure with a center of inversion (Pessen et al., 1991). It was speculated that this asymmetrical structure was due to the formation of a hydrophobically stabilized intermolecular ion pair in α_{s1}-B at this deletion peptide site.

With these two possible dimer structures, a tetramer can easily be modeled starting with the dimer formed by the intermolecular hydrophobic ion pair (Fig. 2.10B). To the two ends of this structure, two molecules of α_{s1}-B are added via the hydrophobic sheet-sheet interaction shown in Fig. 2.10A. Such a structure is presented in Fig. 2.11A. This tetramer structure is highly asymmetric and also contains two possible



hydrophobic ion pair sites at each end of the molecule; these sites at either end of the tetramer could lead to further aggregation resulting in a large rod with a large axial ratio and dipole moment. It is noteworthy that the hydrophobic antiparallel stranded sheets of residues 163–174 could not be docked with the same site on another tetramer structure due to large steric factors. Such a structure could not possibly be made without major changes in the α_{s1}-casein B model. However, a more plausible site for octamer formation via simple hydrophobic interactions can be constructed from use of the hydrophobic side chains centered on prolines 177 and 185, which are located on the lower side of the asymmetric tetramer structure and are solvent accessible. With this in mind, two tetramers were docked in an antiparallel fashion with a center of inversion using these hydrophobic side chains as interaction sites. The octamer structure is shown in Fig. 2.11B. This octamer structure would still allow water to flow through part of the polypeptide chain yielding a high hydrodynamic hydration value. It would also be stable in solution since the hydrophobic side chains are predominately in the center of the model and all eight hydrophilic domains are solvent accessible—two on the

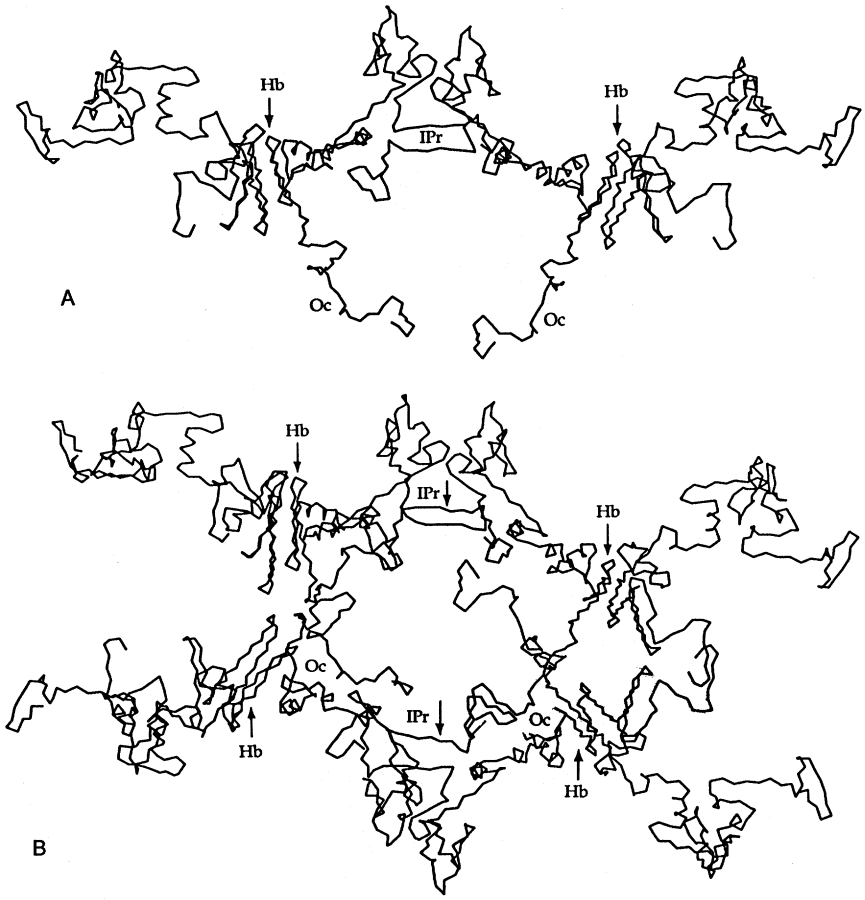


FIG. 2.11 (A) α-Carbon chain trace of α_{s1}-casein B tetramer resulting from the docking of two molecules (left and right) of α_{s1}-B through the hydrophobic (Hb) scheme shown in 2.10A with the (IPr) ion-pair dimer given in Fig. 2.10B. Sites for octamer formation (Oc) are noted. (B) Carbon chain trace of the octamer of α_{s1}-casein B resulting from dimerization of two tetramer structures, shown in 2.11A, via the two hydrophobic (Oc) antiparallel sheet-sheet interaction sites areas related to Hb and IPr sites are also noted.

upper and lower center part and two at either end of the structure (see Fig. 2.11B).

A very interesting feature of all the α_{s1}-caseins is the preservation of the C-terminal tryptophan. Ribabeau-Dumas and Garnier (1970) showed that carboxy-peptidase A could quantitatively remove the C-terminal

tryptophan of α_{s1}-casein alone, and in native and reconstituted micelles. This was interpreted as a demonstration of the open network of the casein micelles that allowed penetration of the protease into the micelle. This is in accord with the model shown in Fig. 2.10A and B, where the C-terminal tryptophan is extended in space at the left side of the model. Thus, although residues 134–185 participate in hydrophobic interactions, a hydrophilic turn then intervenes and the C-terminal tryptophan can still be exposed in monomeric and polymeric structures, making it available to digestion with carboxypeptidase A.

Finally, it can be seen that deletion of residues 14–26 of the α_{s1}-B octamer structure to form α_{s1}-casein A would result in the elimination of the two hydrophobic ion pair–interaction sites. This mutation should result in a tetramer structure arising from only the left- or right-hand portions of the octamer α_{s1}-B structure of Fig. 2.11B.

Thus, we have now established a possible mechanism as well as structures for the calcium-induced salting-out profiles of α_{s1}-casein. Here, divalent cations can bind to serine phosphate or other negatively charged side chains, which are mostly in the hydrophilic domain of α_{s1}-casein (see Fig. 2.9A) causing the hydrophobic domains to interact and self-associate in a cooperative fashion to a limiting octamer structure in the case of α_{s1}-B or a tetramer for α_{s1}-A, respectively. These limiting oligomers are thermodynamically unstable with respect to their interaction with water because of their low charge and solvent exposure to water causing a further noncooperative aggregation leading to precipitation.

Molecular Dynamics of $CaCl_2$ and $MgCl_2$ in Water with the Hydrophilic Domain of α_{s1}-Casein A

In this section we shall attempt to form a structural basis for the thermodynamics of the salting-in process, which was quantitated using thermodynamic linkage and nonlinear regression analysis of the calcium-induced solubility profiles of α_{s1}-casein. Here we have chosen the hydrophilic N-terminal domain of the monomeric α_{s1}-casein A structure. This model was built from the α_{s1}-casein B model of Fig. 2.9A and B by excising residues 14–26 (Eigel et al., 1984) and then deleting residues 100–199 of α_{s1}-B. The resulting hydrophilic domain, i.e., residues 1–99 of the α_{s1}-casein B minus 14–26 (a total of 86 residues), was then energy minimized. Eleven hundred water molecules were added, and the resulting structure in solution was again energy minimized with a cutoff for nonbonded interaction of 5 Å utilizing a periodic boundary condition. The resulting structure with water was subjected to molecular dynamics (MD) calculation for 20 psec above equilibrium conditions, which

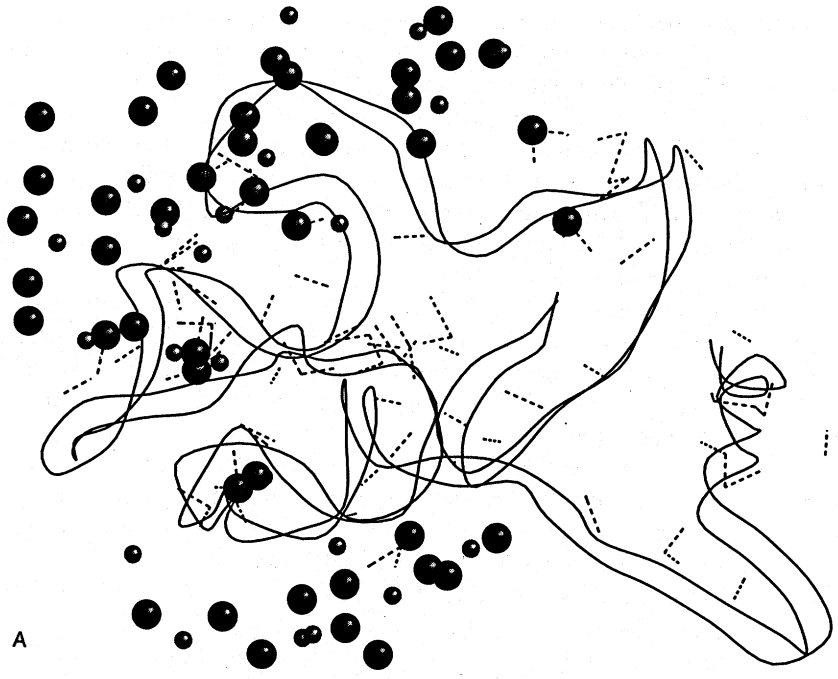


FIG. 2.12 (A) Backbone ribbon structure of the hydrophilic half of α_{s1}-casein A (residue 1 through 86) after molecular dynamics at 40 psec with 22 molecules of $CaCl_2$ in the presence of 1100 water molecules. Ca and Cl atoms are shown as ball models of radii equal to 0.15 times their known van der Waals radius. Dashed lines represent hydrogen bond formation. (B) Stereo view (relaxed) of (A).

were determined by the stabilization of potential energy, radius of gyration of the protein backbone, root-mean-square fluctuations of the backbone atoms, and change in second moment. Such an equilibrated dynamic structure should approximate the structure, energetics, and dynamic motion of this protein domain in solution.

To mimic the salt-binding mechanism, 22 molecules of calcium or magnesium and 44 molecules of chloride with appropriate ionic charges were added to the above system in a pseudo random fashion, energy minimized, and subjected to MD for a full 40 psec. Equilibrium was easily established once again at 15–20 psec. The resulting structure for the native half of α_{s1}-A in $CaCl_2$ is shown in Fig. 2.12A and B. The amount of salt added was chosen to comply with a condition that would result in

saturation of the Ca^{2+}-binding sites. Using a value of 380 M^{-1} for K_A and 8 sites (Dickson and Perkins, 1971), 22 molecules of $CaCl_2$ per 1100 water molecules per molecule of protein is equivalent to greater than 99% occupancy of these calcium-binding sites and greater than 80% occupancy for 8 additional putative salting-in sites derived from k_2 of Table 2.2. A similar structure for the dephosphorylated half of α_{s1}-A was also studied and is shown in Fig. 2.13. In total, seven MD calculations were performed on the hydrophilic domain (H) of α_{s1}-A (residues 1–99) in the presence of 1100 water molecules to 40 psec: two in the absence of salt for native (H) and dephosphorylated (HO-P); two in the presence of $CaCl_2$ for H and HO-P; one with added $MgCl_2$ for H; and one each for $MgCl_2$ and $CaCl_2$ with no protein. Each calculation utilized a cutoff of 5 Å for nonbonded interactions, a Tripos force field, and a "bump" factor of 0.7 for simulating hydrogen bond formation (i.e., a distance of >0.7 Å means no bonding). MD calculations require a running time on the Silicon Graphics Unix computer system of at least 3 days. It should be noted at this time that the energetics and geometric parameters estimated by the MD calculation reflect the total salt binding to the protein, i.e., the sum of the free energies of salt binding for the protein salting-out, as well as the salting-in process. More detailed analysis of these MD calculations, which are beyond the scope of this chapter, must be

FIG. 2.13 Same as in Fig. 2.12A but for dephosphorylated α_{s1}-A. All serine phosphates mutated to serine with appropriate partial charges.

performed for separation of the protein precipitation and resolubilization processes.

The average results of several calculated geometric parameters with their corresponding errors are presented in Table 2.8. Here, the subscript 1 denotes the salt component, i.e., both Ca and Cl, while the subscript 2 describes the protein. R is the calculated dynamic radius of gyration, x is the average center of mass of the component atoms, and a is the root-mean-square fluctuation of component atoms from the center of mass. Here, a can be thought of as a dynamic Stokes radius of the chosen atoms, and x a spherical center of mass.

We now attempt to describe the distribution of the salt atoms at the end of MD calculations by inspection of the a and x values in Table 2.8 for both $CaCl_2$ and $MgCl_2$ alone and in the presence of the hydrophilic half of α_{s1}-A (H) and its dephosphorylated form HO-P. Little significant change is seen in the a_1 values, since the mass of the protein is far greater than that of the salt. The x_1 values are the best descriptors of the binding of salt atoms to the protein in these MD calculations. Here, in all cases the x_1 values have decreased in the presence of either protein component.

TABLE 2.8 Molecular Dynamics of Hydrophilic Half (H) of α_{s1}-casein A in Water with Salt, Geometric Parameters

Protein	Salt	a_1,Å^2	x_1, Å	R_2, Å	a_2, Å^2	x_2, Å
H	—			14.6 ± 0.2	8.2 ± 0.5	0.33 ± 0.04
HO-P	—	—	—	14.2 ± 0.04	7.0 ± 0.2	0.33 ± 0.03
H	$CaCl_2$	11.5 ± 0.7	0.8 ± 0.10	14.6 ± 0.2	8.7 ± 0.4	0.38 ± 0.04
HO-P	$CaCl_2$	9.38 ± 0.67	0.69 ± 0.06	15.5 ± 0.2	10.1 ± 0.4	0.51 ± 0.03
H	$MgCl_2$	7.8 ± 0.6	0.67 ± 0.07	14.6 ± 0.2	7.7 ± 0.5	0.37 ± 0.04
—	$CaCl_2$	9.1 ± 0.9	2.2 ± 0.1	—	—	—
—	$MgCl_2$	6.8 ± 1.1	1.5 ± 0.4	—	—	—

R, radius of gyration; a, RMS fluctuation of all atoms from center of mass (a dynamic Stokes radius); x, calculated spherical center of mass. Subscript 1 denotes salt, 2 denotes protein atoms.

Such a decrease in the average spherical center of mass of the salt atoms is a clear indication of protein-salt interactions, since the center of mass for salt atoms alone would be larger as they move randomly about. However, x_1 would of necessity be smaller for salt bound to protein where movement is restricted. A difference between x_1 for native H and for the O-P form appears, and it can easily be observed that the salt atoms associated with the H and HO-P forms (Fig. 2.12 and 2.13, respectively) have a different overall distribution. This needs to be verified but it is in agreement with the decreased x_1 in the presence of the protein fragments. It is assumed that the x_1 of the $MgCl_2$ was lower due to the lower value of the van der Waals radius for Mg^{+2} (0.66 Å) than for the Ca^{+2} ion (0.99 Å).

To observe the effect of salt binding on the dynamic structure of the hydrophilic half of the protein, we have calculated for the protein in the presence or absence of salt atoms its radius of gyration R_2, as well as the a_2 and x_2 values; these are presented as columns 5, 6, and 7 of Table 2.8. Virtually no changes within the calculated error are observed for the R_2, a_2, or x_2 values for the native H either in the absence or presence of $CaCl_2$ or $MgCl_2$; the HO-P form is not dramatically different either. However, large increases in these descriptors are observed in the dephosphorylated HO-P form when $CaCl_2$ is added in the MD calculations. This change in configuration may reflect a general swelling of the HO-P structure when $CaCl_2$ is added, which can also be observed by inspection of Fig. 2.14A and B. Here the two structures are compared by representations of the protein backbones only, with no side chains displayed. The H form is represented by a ribbon trace of the backbone, while the HO-P form is represented by a backbone–atomic stick model. The backbone model is much more swollen than the ribbon model. For easier observation of this conclusion in all dimensions, a stereo view is shown in Fig. 2.14B.

The reason for this phenomenon is most likely due to the hydrogen bonding of chloride ions to the serine side chains as well as to N-H atoms of the backbone. Such anion-protein hydrogen bonding can impose important dynamic structural changes on the protein component. In the H form these interactions may be "screened out" by the negatively charged phosphate groups. Whether this swelling phenomenon causes increased solubility of the protein, i.e., salting-in, cannot be established at this time, because the deleted residues could play a role in keeping the α_{s1}-B variant insoluble. More MD calculations in conjunction with solution structural physical chemical experiments must be performed in the future to test this hypothesis.

To further correlate the binding free energies calculated from ther-

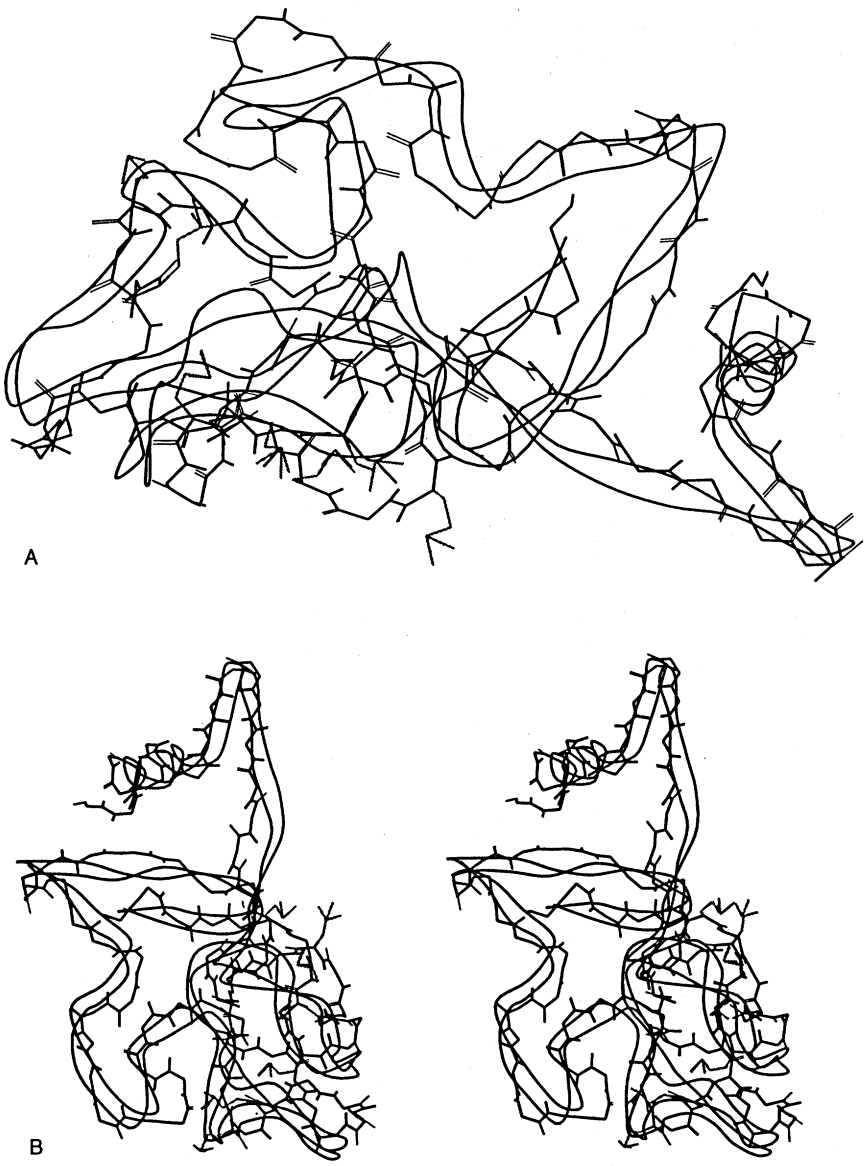


FIG. 2.14 (A) Comparison of hydrophilic domain of α_{s1}-A (shown with ribboned backbone) with dephosphorylated α_{s1}-A after molecular dynamics with 22 molecules of Ca and Cl_2 at 40 psec in the presence of 1100 water molecules. Dashes lines represent hydrogen bond formations. (B) Stereo view (relaxed) of (A).

modynamic linkage analysis of the salting-out and salting-in experiments (Tables 2.5, 2.6), the energetics of the seven MD experiments for the theoretical H and HO-P were calculated. These values are presented in Table 2.9 along with the experimental total binding free energies, ΔF_B, for native α_{s1}-A in $CaCl_2$ and $MgCl_2$ as well as dephosphorylated α_{s1}-A in $CaCl_2$. In the tables, the calculated descriptors of energy are E_T, for the average total potential energy of the system; E_W, for the internal energy of the water; E_{PN}, potential energy of the system; E_W, for the internal energy of the water; E_{PW}, for the energy of the protein, water, and protein-water interaction; E_{SW}, for the salt, water, and salt-water interaction; and E_{int}, for the internal energy for the salt-protein interaction. At this time it should be stressed that the energies estimated from MD calculations are the internal energy at constant volume and not the Gibbs or the Helmholtz free energy derived from binding experiments. However, a qualitative correlation between these two parameters can be utilized to describe variations in protein components or type of salt used. Inspection of column 4 and 5 of Table 2.9 shows that for native α_{s1}-A

TABLE 2.9 Molecular Dynamics of Hydrophilic Half (H) of α_{s1}-casein A in Water with Salt: Energetics and Comparison with Experimental Data

Protein	Salt	$-E_T$	$-E_{int}$	ΔF_B	$-E_{PW}$	$-E_W$	$-E_{SW}$
H	—	20,000	—	—	20,000	13,000	—
		±152			±152	±64	
HO-P	—	19,000	—	—	19,000	13,000	—
		±54			±54	±101	
H	$CaCl_2$	31,000	8,746	33.2	13,000	11,000	20,000
		±134	±162		±154	±142	±128
HO-P	$CaCl_2$	28,000	4,612	14.3	14,000	11,000	22,000
		±132	±198		±26	±91	±155
H	$MgCl_2$	34,000	11,000	36.6	12,000	11,000	22,000
		±92	±175		±133	±142	±155
—	$CaCl_2$	27,000	—		—	13,000	27,000
		±266				±122	±266
—	$MgCl_2$	29,000	—		—	12,000	29,000
		±198				±198	±198

E, potential energy in kcal/mole, i.e., internal energy at constant volume subscripts; *T*, total atoms; *P*, protein atoms; *W*, water atoms; *S*, salt atoms; int, protein-salt interaction. ΔF_B, total cooperative salt-binding free energy in kcal/mole calculated from Tables 2.5 and 2.6 for the solubility data for α_{s1}-casein A (NA) and its dephosphorylated form (NAO-P).

and AO-P in $CaCl_2$ and native A in $MgCl_2$, the experimental changes of ΔF_B correlate well with changes in ΔE_{int} when these environmental factors are varied; the absolute values for the two parameters are, however, significantly different. These large differences in absolute values between ΔF_B and ΔE_{int} are due to the bump factor of 0.7 used in these calculations to quantitate hydrogen bonding in MD. If this factor were optimized then perhaps better correlations could be achieved. However, for this study, the value of 0.7 was used since it is the default value used by Sybyl for the Tripos force field. Such a small bump factor most likely allows for inordinately high electrostatic energy terms. Future work will tend to optimize this parameter. No conclusions can be made at this time concerning the meaning of the other energetic descriptors of Table 2.9, but they are presented along with their error for inspection by the reader.

SUMMARY AND CONCLUSIONS

It can be concluded that thermodynamic linkage in conjunction with molecular modeling techniques such as energy minimization and molecular dynamics can provide a powerful multifaceted approach for developing structure-function relationships such as the salt-induced solubility profiles of this study. Here, the stoichiometry and thermodynamics of the salt-induced solubility can be quantitated using thermodynamic linkage in conjunction with nonlinear regression analysis. The dynamic changes in the protein structure responsible for the salting-out and salting-in processes can be established using energy minimization and molecular dynamics calculations of the protein in water and in the presence and absence of salt atoms. In particular, predictions concerning the type and amount of protein modification that occurs can be utilized to increase the desired protein functionality in a rational way through chemical or genetic modification (Richardson et al., 1992).

For example, we have structurally defined the salting-out process in terms of a hydrophobically controlled octamerization for α_{s1}-B in compliance with the experimental results of Waugh and et al. (1971) and our quantitation of their results using thermodynamic linkage. This oligomerization is most likely caused by binding of the positive divalent cations to the phosphoserine and other negatively charged side chains on the hydrophilic domain of α_{s1}-casein (see Fig. 2.9A) resulting in the minimization of protein charge and the protein-protein electrostatic repulsion term. The first step in this self-association is most probably caused by the dimerization of both α_{s1}-B and α_{s1}-A, which proceeds by

the docking of two large hydrophobic stranded antiparallel sheets in an asymmetric fashion (see Fig. 2.10A). Disruption of this interaction via modification of these large antiparallel sheets by chemical, genetic, or biochemical means may significantly change the solubility of α_{s1}-B or -A in the presence of $CaCl_2$. An example of this could be plasmin cleavage of α_{s1}-B to yield two large peptides (H. E. Swaisgood, personal communication). This potential enzymatic cleavage site (residues 101–104) is between the hydrophobic and hydrophilic domains of α_{s1}-casein (see Fig. 2.9A). The salting-in process appears to be more complicated; however, it can be naively described in a general fashion as further salt binding by the protein's hydrophilic domain allowing resolubilization by increased protein charge, resulting in a large protein-protein electrostatic repulsion term. The difference in the calcium-induced salting-in between the α_{s1}-B variant and the α_{s1}-A variant was structurally defined in terms of an intermolecular hydrophobically stabilized ion pair (see Fig. 2.10B) which is present in α_{s1}-B and which disrupts the resolubilization process. The deletion of residues 16–24 of α_{s1}-B results in the α_{s1}-A variant, which is more soluble with added calcium. The α_{s1}-A genetic variant has a gene frequency of 0.01% and is, therefore, not present in most cows. Hence, the possible enzyme cleavage of residues 1–24 of α_{s1}-B or chemical modification may aid in making this variant more soluble with added $CaCl_2$.

Finally, the molecular dynamics calculations could not at the present time distinguish between the salting-out and salting-in binding free energy. Only the total binding could be correlated in the MD results. In fact, significant differences in absolute values occur when the total salt-binding free energy calculated from thermodynamic linkage, ΔF_B, is compared with the interaction energy derived from MD, ΔE_{INT} (see Table 2.8). This discrepancy is most likely due to the bump factor value chosen for the MD calculations; a better value could be chosen for future works. However, the changes in ΔF_B with the change in salt from Ca to Mg for the native systems as well as the change with calcium between native and the dephosphorylated form show good correlations with ΔE_{INT}. More importantly, the geometric parameters for the protein from MD calculations, i.e., R_2, a_2, and x_2 of the HO-P, change dramatically over the H form in the presence of $CaCl_2$. This phenomenon was interpreted as a swelling of the dephosphorylated form with added $CaCl_2$ (Fig. 2.14A and B). It is apparent that this structural change can also be labeled as a conformational change since many internal hydrogen bonds are disrupted. This could be tested experimentally by FTIR.

However, this general, swelling of the HO-P form with added $CaCl_2$ can lead to large thermal fluctuation as seen by the a_2 parameter of Table

2.8. This increase in motion would lead to an increase in the entropy of mixing and ultimately would increase the solubility of the protein. To what extent this entropy of mixing is important to the resolubilization process cannot be assessed at the present time. These large fluctuations of HO-P with added $CaCl_2$ may in the future also aid in increasing other desired functionalities such as gelation, whippability, or foam formation.

REFERENCES

Andersen, H. C. 1980. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.* 72: 2384–2394.

Arakawa, T. and Timasheff, S. N. 1984. Mechanisms of protein salting-in and salting-out by divalent cation salts: balance between hydration and salt binding. *Biochemistry* 23: 5912–5923.

Bingham, E. W., Farrell, H. M., Jr., and Carroll, R. J. 1972. Properties of dephosphorylated α_{s1}-casein. Precipitation by calcium ions and micelle formation. *Biochemistry* 11: 2450–2454.

Cann, J. R. 1978. Measurements of protein interactions mediated by small molecules using sedimentation velocity, In *Methods in Enzymology XLVIII*, C. H. W. Hirs and S. N. Timasheff (Ed.), pp. 242–248. Academic Press, NY.

Cann, J. R., and Hinman, N. D. 1976. Hummel-Dryer gel chromatographic procedures as applied to ligand-mediated associations. *Biochemistry* 15: 4614–4628.

Creamer, L. K., and Waugh, D. F. 1965. Calcium binding and precipitate solvation of Ca-α_s-caseinates. *J. Dairy Sci.* 49: 706.

Dickson, I. R., and Perkins, J. D. 1971. Studies on the interactions between purified bovine caseins and alkaline earth metal ions. *Biochem. J.* 124: 235–240.

Eigel, W. N., Butler, J. E., Ernstrom, C. A., Farrell, H. M., Jr., Harwalkar, V. R., Jenness, R., and Whitney, R. McL. 1984. Nomenclature of the proteins of cows' milk: 5th Revision. *J. Diary Sci.* 67: 1599–1631.

Farrell, H. M., Jr., and Kumosinski, T. F. 1988. Modeling of calcium-induced solubility profiles of casein for biotechnology. *J. Industrial Micro.* 3: 61–71.

Farrell, H. M., Jr., and Thompson, M. P. 1988. The caseins of milk as calcium binding proteins. In *Calcium Binding Proteins*, M. P. Thompson (Ed.), pp. 117–137. CRC Press, Boca Raton, FL.

Farrell, H. M., Jr., Kumosinski, T. F., Pulaski, P., and Thompson, M. P. 1988. Calcium-induced associations of the caseins: A thermodynamic

linkage approach to precipitation and resolubilization. *Arch. Biochem. Biophys.* 265: 146–158.

Kollman, P. A. 1987. Application of force fields to molecular models. *Ann. Review Phys. Chem.* 38: 303–333.

Kumosinski, T. F., and Farrell, H. M., Jr. 1991. Calcium-induced associations of the caseins: Thermodynamic linkage of colloidal stability of casein micelles to calcium binding. *J. Protein. Chem.* 10: 3–16.

Kumosinski, T. F., Brown, E. M., and Farrell, H. M., Jr. 1991a. Molecular modeling in food research: Technology and techniques. *Trends Food Sci. Technol.* 2: 110–115.

Kumosinski, T. F., Brown, E. M., and Farrell, H. M., Jr. 1991b. Molecular modeling in food research: Applications. *Trends Food Sci. Technol.* 2: 190–195.

Kumosinski, T. F., Brown, E. M., and Farrell, H. M., Jr. 1993. Three dimensional molecular modeling of bovine caseins: An energy-minimized β-casein structure. *J. Dairy Sci.* 76: 931–945.

Kumosinski, T. F., Brown, E. M., and Farrell, H. M., Jr. 1994. Three dimensional molecular modeling of bovine caseins: an energy minimized α_{s1}-casein structure. In *Molecular Modeling,* T. F. Kumosinski and H. N. Liebman (Eds.), ACS Symposium Series, Denver, CO.

Melander, W. and Horvath, C. 1977. Salt effects on hydrophobic interactions in precipitation and chromatography of proteins: An interpretation of the lyotropic series. *Arch. Biochem. Biophys.* 183: 200–215.

Noble, R. W. and Waugh, D. F. 1965. Casein micelles, formation and structure I. *J. Am. Chem. Soc.* 87: 2236–2245.

Pessen H., Kumosinski, T. F., Farrell, H. M., Jr., and Brumberger, H. 1991. Tertiary and quaternary structural differences between two genetic variants of bovine casein by small-angle X-ray scattering. *Arch. Biochem. Biophys.* 284: 133–142.

Ribadeau-Dumas, B. and Garnier, J. 1970. Structure of casein micelle. The accessability of subunits to various reagents. *J. Dairy Res.* 37: 269–278.

Richardson, T., Jiminez-Flores, R., Kumosinski, T. F., Oh, S., Brown, E. M., and Farrell, H. M., Jr. 1992. Molecular modeling and genetic engineering of milk proteins. In *Advanced Diary Chemistry 1: Proteins,* P. F. Fox (Ed.), pp. 545–578. Elsevier, Essex, UK.

Robinson, D. R., and Jencks, W. P. 1965. Effects of concentrated salt solutions on the activity coefficient of acetyltetraglycine ethylester. *J. Am. Chem. Soc.* 87: 2470–2479.

Schmidt, D. G. 1982. Association of caseins and casein micelle structure. In *Developments in Diary Chemistry*, P. F. Fox (Ed.), pp. 61–86. Applied Science Publications Ltd., London, UK.

Sillen, L. G. and Martell, A. E. 1971. *Stability Constants of Metal Ion Complexes.* Special Publication No. 25 of The Chemical Society of London. Alden Press, Oxford, UK.

Sinanoglu, O. 1968. Solvent effects on molecular association. In *Molecular Associations in Biology*, B. Pullman (Ed.), pp. 429–445. Academic Press, New York.

Tanford, C. 1967. *Physical Chemistry of Macromolecules.* John Wiley & Sons, New York.

Thompson, M. P., Gordon, W. G., Boswell, R. T., and Farrell, H. M., Jr. 1969. Solubility, solvation and stabilization of α_{s1}- and β-caseins. *J. Dairy Sci.* 52: 1166–1173.

Waugh, D. F. and Noble, R. W. 1965. Casein micelles, formation and structure II. *J. Am. Chem. Soc.* 84: 2246–2257.

van Gunsteren, W. F. and Berendsen, H. J. C. 1977. Algorithms for molecular dynamics and constraint dynamics. *Mol. Phys.* 34: 1311–1327.

Waugh, D. F., Slattery, C. W., and Creamer, L. K. 1971. Binding of calcium to caseins. *Biochemistry* 10: 817–823.

Weiner, S. J., Kollman, P. A., Nguyen, D. T., and Case, D. A. 1986. Force field calculations in computational chemistry. *J. Comput. Chem.* 7: 200–230.

Wyman, J., Jr. 1964. Linked functions and reciprocal effects in hemoglobin: A second look. *Adv. Protein Chem.* 19: 223–286.